Medical Physics
The International Journal of Medical Physics Research and Practice

Impact of compressed breast thickness and dose on lesion detectability in digital mammography: FROC study with simulated lesions in real mammograms

Elena Salvagnini, Hilde Bosmans, Chantal Van Ongeval, Andreas Van Steen, Koen Michielsen, Lesley Cockmartin, Lara Struelens, and Nicholas W. Marshall

Citation: Medical Physics **43**, 5104 (2016); doi: 10.1118/1.4960630
View online: http://dx.doi.org/10.1118/1.4960630
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/43/9?ver=pdfcov
Published by the American Association of Physicists in Medicine

Impact of compressed breast thickness and dose on lesion detectability in digital mammography: FROC study with simulated lesions in real mammograms

Elena Salvagnini[a)]
Department of Imaging and Pathology, Radiology, KUL, Herestraat 49, Leuven B-3000, Belgium and SCK•CEN, Boeretang 200, Mol 2400, Belgium

Hilde Bosmans
Department of Imaging and Pathology, Radiology, KUL, Herestraat 49, Leuven B-3000, Belgium and Department of Radiology, Radiology, UZ Gasthuisberg, Herestraat 49, Leuven B-3000, Belgium

Chantal Van Ongeval and Andreas Van Steen
Department of Radiology, Radiology, UZ Gasthuisberg, Herestraat 49, Leuven B-3000, Belgium

Koen Michielsen
Department of Imaging and Pathology, Nuclear Medicine and Molecular Imaging, KUL, Herestraat 49, Leuven B-3000, Belgium

Lesley Cockmartin
Department of Radiology, Radiology, UZ Gasthuisberg, Herestraat 49, Leuven B-3000, Belgium

Lara Struelens
SCK•CEN, Boeretang 200, Mol 2400, Belgium

Nicholas W. Marshall
Department of Imaging and Pathology, Radiology, KUL, Herestraat 49, Leuven B-3000, Belgium and Department of Radiology, Radiology, UZ Gasthuisberg, Herestraat 49, Leuven B-3000, Belgium

(Received 19 January 2016; revised 20 July 2016; accepted for publication 27 July 2016; published 16 August 2016)

Purpose: The aim of this work was twofold: (1) to examine whether, with standard automatic exposure control (AEC) settings that maintain pixel values in the detector constant, lesion detectability in clinical images decreases as a function of breast thickness and (2) to verify whether a new AEC setup can increase lesion detectability at larger breast thicknesses.
Methods: Screening patient images, acquired on two identical digital mammography systems, were collected over a period of 2 yr. Mammograms were acquired under standard AEC conditions (part 1) and subsequently with a new AEC setup (part 2), programmed to use the standard AEC settings for compressed breast thicknesses ≤49 mm, while a relative dose increase was applied above this thickness. The images were divided into four thickness groups: T1 ≤ 29 mm, T2 = 30–49 mm, T3 = 50–69 mm, and T4 ≥ 70 mm, with each thickness group containing 130 randomly selected craniocaudal lesion-free images. Two measures of density were obtained for every image: a BI-RADS score and a map of volumetric breast density created with a software application (VolparaDensity, Matakina, NZ). This information was used to select subsets of four images, containing one image from each thickness group, matched to a (global) BI-RADS score and containing a region with the same (local) VOLPARA volumetric density value. One selected lesion (a microcalcification cluster or a mass) was simulated into each of the four images. This process was repeated so that, for a given thickness group, half the images contained a single lesion and half were lesion-free. The lesion templates created and inserted in groups T3 and T4 for the first part of the study were then inserted into the images of thickness groups T3 and T4 acquired with higher dose settings. Finally, all images were visualized using the ViewDEX software and scored by four radiologists performing a free search study. A statistical jackknife-alternative free-response receiver operating characteristic analysis was applied.
Results: For part 1, the alternative free-response receiver operating characteristic curves for the four readers were 0.80, 0.65, 0.55 and 0.56 in going from T1 to T4, indicating a decrease in detectability with increasing breast thickness. *P*-values and the 95% confidence interval showed no significant difference for the T3-T4 comparison ($p = 0.78$) while all the other differences were significant ($p < 0.05$). Separate analysis of microcalcification clusters presented the same results while for mass detection, the only significant difference came when comparing T1 to the other thickness groups. Comparing the scores of part 1 and part 2, results for the T3 group acquired with the new AEC setup and T3 group at standard AEC doses were significantly different ($p = 0.0004$), indicating improved

detection. For this group a subanalysis for microcalcification detection gave the same results while no significant difference was found for mass detection.
Conclusions: These data using clinical images confirm results found in simple QA tests for many mammography systems that detectability falls as breast thickness increases. Results obtained with the AEC setup for constant detectability above 49 mm showed an increase in lesion detection with compressed breast thickness, bringing detectability of lesions to the same level. © *2016 American Association of Physicists in Medicine.* [http://dx.doi.org/10.1118/1.4960630]

Key words: Digital mammography, lesions detectability, FROC, lesions simulation

1. INTRODUCTION

In many countries, full field digital mammography (FFDM) has largely replaced screen-film mammography for the purpose of screening program and diagnostic follow up. One of the consequences of the large scale implementation of mammography has been a thorough investigation of many aspects of the mammography system setup together with the optimization of system subcomponents. Several studies have compared the detectability performance of FFDM against screen-film mammography in clinical practice[1–7] and found FFDM to be equal or noninferior to screen-film mammography. In terms of FFDM system optimization, investigators have looked at beam quality, detector, system efficiency, and operating dose level. Regarding beam quality, anode and filter settings have been explored that maximized image quality at the lowest dose.[8–11] Other studies have investigated the efficiency of mammographic systems in term of effective detective quantum efficiency or effective noise equivalent quanta.[12,13] Potential for dose reduction has been explored using simulated lesions.[14,15]

Notwithstanding the extensive investigation of FFDM system setup in the literature, there has been little work on lesion detectability for the range of different breast thicknesses seen in clinical practice. Recent studies have shown that theoretical detectability/visibility indexes such as signal difference to noise ratio (SDNR) and threshold contrast-detail (*c-d*) detectability decrease when test object thickness increases,[16–19] leading to a potential loss of lesion visibility in clinical practice. This is a consequence of the automatic exposure control (AEC) programming for many of the mammographic systems on the market in which AEC devices are designed or configured to maintain the signal, derived from the image receptor, constant for all beam qualities, irrespective of breast composition or thickness. This work examines whether, for the standard AEC settings, simulated lesion detectability decreases as a function of breast thickness and whether new AEC settings that keep technical image quality metrics constant as a function of thickness would bring lesion detectability to the same level for all breast thicknesses. This is achieved in a two-part study using simulated lesions (both microcalcifications and masses), scored by experienced radiologists in a free-response observer study. Images for the first part of the study were acquired using the standard AEC setting. New AEC settings were then programmed and used to acquire the images used in the second part of the study. An in-depth description and derivation of the new AEC has been given in a previous study,[16,17] along with detailed validation using the CDMAM (Artinis BV, The Netherlands) *c-d* test object as a function of poly(methyl methacrylate) (PMMA) thickness. Hence, the aim of this study is to extend the previous theoretical/phantom based work with a validation of the new AEC setup using real mammograms.

2. MATERIAL AND METHODS

2.A. Selection of mammograms

Image acquisition was as follows. Patient images were collected during two years of routine screening examinations performed at the University Hospital of Leuven, Belgium. During this period, both "for processing" and "for presentation" images, acquired on two Siemens Mammomat Inspiration systems (Siemens, Erlangen, Germany), were collected and stored. The first year's images were acquired using the system standard automatic exposure control settings. During the second year, the automatic exposure control was set to use higher tube current-time product (mAs) values, and hence radiation doses, to obtain technical detectability indexes identical for breast equivalent thicknesses above 49 mm. The methodology behind the modified AEC scheme is described in the study of Salvagnini *et al.*[16,17] which showed that the AEC can be set up to keep detectability constant for test object thicknesses above 40 mm, while maintaining mean glandular doses (MGDs) below both the Acceptable and Achievable levels of the European Guidelines.[20]

Following image acquisition, the simulation stage began. The complete image database was analysed with the VOLPARA volumetric density software (VolparaSolutions, Wellington, New Zealand)[21,22] and a glandularity map was created for each mammogram. Although VOLPARA uses a proprietary algorithm, the basic principle of operation is outlined in a number of publications.[21,23] An area of the breast that is composed entirely of fatty tissue is required/found and the pixel value (PV) within this region is taken as the reference level from which the thickness of dense tissue overlying every pixel in the image can be estimated. Additional information required in making this estimate is the effective x-ray linear attenuation coefficients for fatty and dense tissues for the breast thickness and x-ray spectrum used.[21,23] In part one, 520 craniocaudal mammograms were selected from patient images acquired during the first year of data collection. The inclusion criterion was the absence of lesions, apart from

calcified vessels and macrocalcifications, considered to hold no relevance for the subsequent reading studies. During the selection of the images, the BI-RADS density score was assigned to each mammogram by an expert radiologist with more than 20 yr of experience in breast cancer imaging. Four groups, each with 130 images, were created according to compressed breast thickness: group T1 ≤ 29 mm, group T2 = from 30 to 49 mm, group T3 = from 50 to 69 mm, group *T*4 ≥ 70 mm. The images were used as normal backgrounds for the insertion of simulated masses and microcalcification clusters. In the second part of the study, 260 craniocaudal mammograms were selected from patient images acquired during the second year of data collection with the modified AEC setting. Again, 130 mammograms were selected for inclusion in thickness groups T3 and T4, following the same criteria as in part 1. The BI-RADS score was also assigned by the same expert radiologist. For the second part of the study, only thickness groups T3 and T4 were investigated; the reason for this is outlined in the preceding study that described the theory behind the new AEC settings.[16] In that work, an AEC design was discussed that held detectability of targets at some constant level (i.e., the Achievable image quality level in the EC guidelines) as breast thickness changed. While this setting would actually reduce the dose to thinner breasts (i.e., breasts in thickness groups T1 and T2), it is likely that the dose reduction would be accompanied by a reduction in image quality and hence the decision was taken not to apply the new AEC setup at small breast thickness. This led to the idea of a composite AEC where the original AEC function (that held PV constant) was left in place for groups T1 and T2 and the new AEC control, that increased detector PV, was applied for thicker breasts (groups T3 and T3). The main idea was therefore to improve image quality and lesion detectability at large breast thicknesses and to try and bring this closer to that for small breast thicknesses, while maintaining breast doses below the European Guideline limits for digital mammography. Hence, for the second part of the study, the setup of the AEC was a combination of the standard AEC mode (patient breast thickness ≤ 49 mm) and the new AEC mode (patient breast thickness ≥ 50 mm). Average MGD for the two parts was calculated from the organ dose values held in the DICOM headers using tag {0040,0316} "Organ Dose"; accuracy of the stored organ dose value for these two systems is checked during biannual QC checks and was within 5% of the calibrated value.

2.B. Simulated lesions and simulation tool

A database of 70 lesions, of which 35 were simulated microcalcification clusters and 35 were simulated masses, was used in this study. The creation of the lesion models used in this study is described in the work of Shaheen *et al.*,[24,25] along with their validation in terms of realism in FFDM and in breast tomosynthesis. Both spiculated and nonspiculated masses were used.[24] Similar models of simulated microcalcification clusters had been validated and used in previous observer studies to compare image processing algorithms.[26,27] Simulated lesions rather than real pathological cases were chosen for this study as this allowed control over the number of lesions in the different thickness groups along with the contrast, position, and size of the lesions, without having to acquire thousands of screening mammograms especially for the quite sparsely populated subgroups like very thick or very thin breasts. The possibility of controlling and fixing these parameters is fundamental to the investigation of the influence of breast thickness on lesion detectability.

The creation of abnormal cases was performed by inserting simulated lesions in normal mammograms using a simulation tool developed to simulate 3D objects into 2D mammograms and validated previously by Shaheen *et al.*, a detailed description of which can be found elsewhere.[24,28] The simulation accounts for the geometry of the mammographic system and for system specific characteristics such as presampling modulation transfer function (MTF) and scatter fraction (SF). The scatter values used in the simulation tool for this study were taken from the work of Salvagnini *et al.*[29] The lesion template was also adjusted for clinical parameters specific to the mammogram of interest, such as compressed breast thickness, acquisition spectrum, background glandularity percentage and beam quality. Moreover, the composition and diameter of the simulated lesion can be specified. In this work the materials selected for the simulation of masses and microcalcifications were 100% glandular tissue and calcium oxalate, respectively. The minimum and the maximum diameters of the simulated microcalcification clusters were 1 and 6.2 mm. The diameter of the largest microcalcification present in the clusters ranged from 0.31 to 1.46 mm. The average diameter of the largest microcalcification of the study was 0.65 mm. For the simulated masses, the diameters ranged from 4.6 to 16.8 mm. These diameters were calculated from the template of the lesion before insertion in the mammogram. Lesion templates were inserted into for processing DICOM images and then processed.

2.C. Creation of abnormal cases

Each thickness group contained 130 images of which 60 were used as normal images and 70 were used to create abnormal cases; in half of these images simulated microcalcification cluster (35 different models) were inserted; in the other half, simulated masses (35 different models) were inserted. Lesions were always simulated into for processing images.

To perform the insertion of a lesion, four patient images were selected at the same time, namely one image for each thickness group. Two specific criteria were required for the selection of these images: (1) the same BI-RADS density score and (2) an area of the image with the same glandularity percentage (g%). The glandularity percentage values were taken from VOLPARA glandularity maps. An example of four such mammograms, with BI-RADS density score equal to 2, is shown in Fig. 1.

The selection criteria for the images play an important role in controlling the influence of mammographic background structure on lesion detection and isolating the effect of thickness. Lesion templates were created to be "subtle," "not

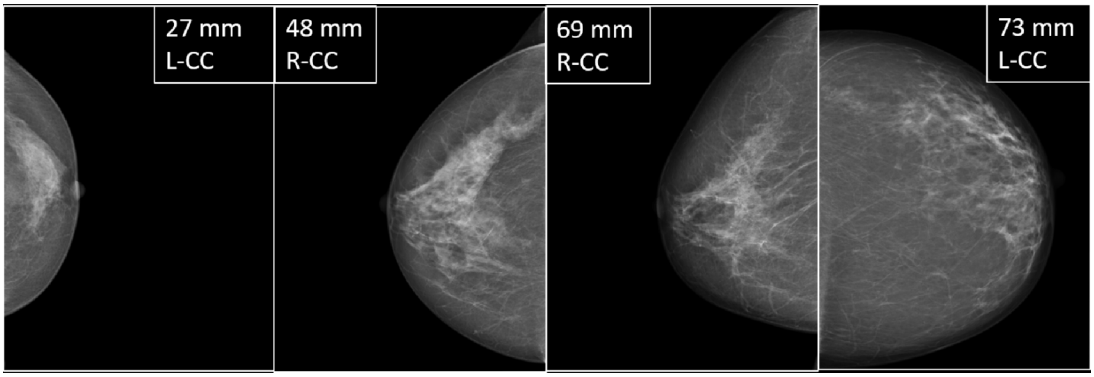

FIG. 1. Four mammograms of four different thickness groups, presenting the same BI-RADS score. These mammograms were selected for the insertion of the same template.

visible" or "obvious" in mammograms having compressed breast thickness between 40 and 59 mm, i.e., group T2. This thickness range was selected because the average breast thickness of the female population falls in this range. In particular for the subtle cases, 17 masses and 17 microcalcification clusters were created in patients with compressed breast thickness between 40 and 49 mm while 7 masses and 8 microcalcification clusters were created in mammograms with thickness ranging between 50 and 59 mm. Creation of a lesion template entailed adjusting the 3D model in size and thickness until it could be considered to be subtle, obvious, or not visible after processing the image. Appearance and positions of the templates were verified by an experienced radiologist who did not take part in the reading study. After the creation of a template for one specific thickness group, the same model was inserted with identical attenuation coefficient and adjusted scatter fraction in the other 3 thickness groups in images presenting the same BI-RADS score and local percentage glandularity score. Examples of zoomed areas of 21% local glandularity value before and after the insertion of a mass are shown in Fig. 2. These image segments were taken from the mammograms shown in Fig. 1.

In part 2 of the study, the same lesion templates created and used in part 1 were then inserted into the new set of mammograms acquired with the new AEC setup. As explained in Sec. 2.A, the new AEC settings only affected breasts with thicknesses falling into groups T3 and T4. Images in which a lesion was simulated were matched in terms of the BI-RADS density score, local glandularity score, and compressed breast thickness of these mammograms with the mammograms already used in part 1. Keeping these three parameters equal between the selected mammograms allowed the influence of dose (and hence the new AEC setting) on lesion detectability to be isolated while minimizing the effects

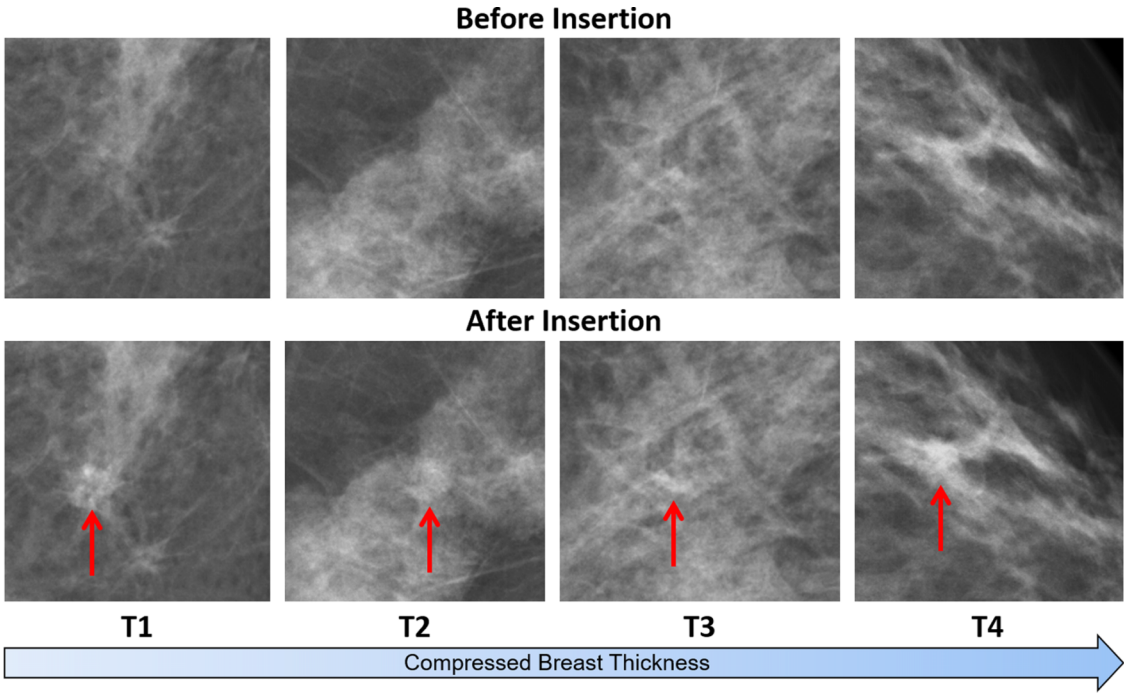

FIG. 2. Zoom of the areas selected for insertion from the mammograms in Fig. 1. The areas had a g% value equal to 21%. Areas before lesion insertion are shown in the upper row and the same areas after lesion insertion are represented in the lower row.

of background glandularity and breast thickness. For both parts of the study, lesions were simulated in the whole breast area (x,y) but avoiding regions of the pectoral muscle, the skin edge, and the nipple. In terms of depth within the breast (z position), lesions were always placed half way between the compression paddle and the detector cover. The local glandularity percentage values of the backgrounds ranged from 7% to 28% for masses and from 7% to 31% for microcalcification clusters. Each abnormal case contained only one simulated lesion. All images, normal and abnormal, were processed with the Siemens standard image processing algorithm for 2D digital mammography, OpView2 (Siemens, Erlangen, Germany).

2.D. Observer performance study

Four radiologists participated in the study. They had between 1 and 20 yr of experience in reading mammograms and in particular between 1 and 10 yr of experience with digital mammography images; they read at least 20 000 mammograms/yr. The observer performance study consisted in a free-response paradigm.[30,31] All observers read all images of both study parts and were told that each case could contain zero, one or more lesions. Readers were asked to mark the center of any region where they suspected the presence of a lesion, be it a mass or a microcalcification cluster. Image display and flow was controlled using ViewDEX software,[32,33] which allows the user to pan, zoom, window/level and put marks in an image. After a mark has been placed, a small cross remains on the image. Marks can be hidden or deleted during the scoring of the image; it is not possible to change the score of a previous image. Following the mark up of a lesion, the observer was presented with three questions:

- What is the confidence of the presence of the lesion? This question was answered on a 5-point scale: (1) Very unlikely to be a lesion. (2) Probably not a lesion. (3) Possibly a lesion. (4) Probably a lesion. (5) Definitely a lesion.
- What is the lesion type? Observers could choose between mass or microcalcification cluster.
- What is the likelihood that this lesion is malignant? This question was also answered on a 5-point scale: (1) Definitely benign. (2) Probably benign. (3) Possibly malignant. (4) Probably malignant. (5) Definitely malignant. (Note that results of this question are not used for in this study, and therefore not reported.)

The software recorded each score of an observer with the relative coordinates, the answers to the questions, and image reading time.

The first part of the study consisted of 13 sets of images each containing 40 images; the second part of the study consisted of 7 image sets, 6 sets of 40 images, and one of 20 images. An extra section of 24 images was for training before the actual study commenced. Average reading time per session was 35 min. All images were displayed on 5 megapixel medical grade Barco monitors (Barco, Kortrijk, Belgium) which were calibrated to the DICOM grayscale standard display function GSDF. Ambient light was <10 Cdm^{-2} and the monitors were positioned such that screen reflections were eliminated, a situation that represents the standard clinical conditions for mammography reading.

2.E. Statistical analysis

A lesion can be marked or not marked, leading to the standard range of false positive (FP), true positive (TP), false negative (FN), and true negative (TN) results. To determine the category of each mark, an acceptance radius of 25 pixels (2.125 mm) was defined around the lesion center to the mark. If the mark was placed within this radius, the result was considered as a TP.[34] This classification of the marks together with their rating was used as input file for the jackknife-alternative free-response receiver operating characteristic (JAFROC) software (JAFROC, version 4.2.1, Chakraborty) available from http://www.devchakraborty.com. This software was used to perform a JAFROC analysis.[30] In this analysis, the figure of merit (FoM) is the probability that a lesion rating exceeds the highest rated nonlesion rating on normal images and is calculated as the trapezoidal area under the alternative free-response receiver operating characteristics (AFROC). The software performs significance testing using the ANOVA technique described by Dorfman, Berbaum, and Metz.[35,36] The software produced different results when analyzing the cases and the readers as random or fixed. In this study, only results for random cases and random readers will be presented because these results can be applied to the whole population of readers and cases. The overall results for the significance of the statistical analyses are presented as F-statistics and degrees of freedom for both parts of the study. The statistical significance for comparison of the modalities, intended as thickness groups for part 1 of the study and as doses for part 2 of the study, is presented as the AFROC area, p-values, and the 95% confidence interval.

3. RESULTS

3.A. Mean glandular dose and exposure time

Before discussing the observer results in detail we first give some breast dose and exposure time information for the standard and modified AEC modes. For the part 1 dataset, average MGDs for thickness groups T1 (≤29 mm), T2 (30–49 mm), T3 (50–69 mm), and T4 (≥70 mm) were 0.90, 1.03, 1.20, and 1.57 mGy, respectively. Following the dose increase for part 2, the average MGD was 1.94 mGy for group T3 and 2.67 mGy for group T4, indicating an increase of dose of 60% and 70% for the two investigated thickness groups. Despite the higher MGDs given to patients in part 2, doses remained below the Achievable dose limit indicated in the European guidelines.[20] The mean glandular doses of the patients used in parts 1 and 2 of the study are presented in Fig. 3 for the different thickness groups. Because patients were chosen for their thickness only, the distribution of BI-RADS scores is random and includes all categories.

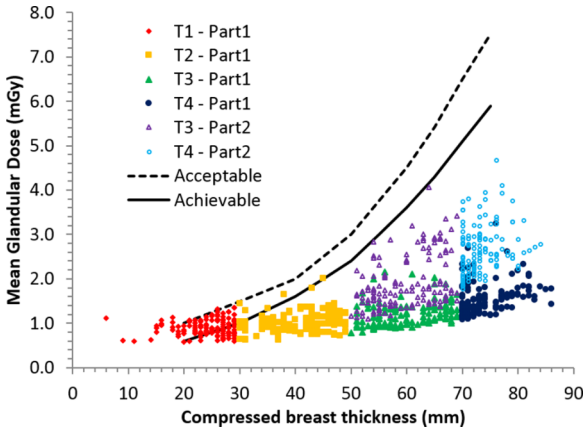

FIG. 3. Patient MGD for parts 1 and 2 of the study, divided by thickness group. The Acceptable and the Achievable dose limits prescribed by the European guidelines are also indicated.

The patient doses for part 2 of the study for T3 and T4 thickness groups shown in Fig. 3 have a larger spread compared to the same thickness groups for the standard AEC mode. This is probably due to the modified AEC aiming for a higher target pixel value for the thicker breasts. This target PV for the modified AEC mode is 60% and 70% higher for T3 and T4 respectively, resulting in a larger range of mAs values that can be selected in reaching the target. The AEC selects a certain mAs value based on a preshot exposure that searches the most dense area of the breast. Local breast density variation across the breast will produce a larger variation of mAs values and hence an increased range of MGD values for the women.

Accompanying the dose increase is a lengthening of exposure time which can increase the probability of motion blurring and ultimately lead to blurring of the imaged anatomy. In this study the average exposure times in part 1 for the different thickness groups were 0.56, 0.59, 0.71, and 0.98 s, for groups T1 to T4, respectively. For part 2 the average exposure times were 1.11 and 1.63 s for T3 and T4 respectively, an increase of 56% and 66% over part 1. This increase of exposure time follows that of the MGD increase, as expected if tube current is held constant. The exposure time values found for part 2 were compared to those in the work of Berns *et al.*[37] for screen-film and digital mammography, where exposure times for 4.5 and 6 cm of breast equivalent material were respectively 1.05 and 1.61 s, for screen-film mammography and 0.70 and 0.79 s for digital mammography. Hence average exposure times for study for part 2 were similar to those for screen-film mammography, with average breast thickness equal to 59.6 and 73.3 mm, for groups T3 and T4, respectively.

TABLE II. Shown is the overall significance of the test for the different parts of the study. The numerator and the denominator of the degrees of freedom, the F-statistics, and the p-value are presented. A p-value smaller than 0.05 means that at least one pair of modalities compared in the test is significantly different. The symbol ˙ indicates no significant difference.

Study	Degree of freedom (numerator, denominator) F-statistics	p-value
Part 1	(3, 113.99) 16.39	<0.0001
Part 2 (50–69 mm)	(1, 219.17) 12.79	0.0004
Part 2 (>70 mm)	(1, 6.55) 2.29	0.1305˙

3.B. Observer detection results

Table I presents the number of correct lesion marks (TP) for the different thickness groups. Each cell shows the number of the lesions marked correctly together with the total number of lesions present in the specific thickness groups and the percentage of correct lesion marks. Columns 2–4 show the results of part 1 of the study in which the influence of compressed breast thickness on lesion detectability is investigated for the system standard AEC setup. The results of part 2 of the study, using the modified dose setting, are shown in columns 5–7. Results for both parts are shown for all lesions, and for microcalcification clusters and masses separately.

Table II presents the overall results of the JAFROC analysis for the different parts of the study. In this table the degrees of freedom and the F-statistic together with the p-value are shown. A p-value inferior to 0.05 means that at least one pair of modalities compared in the test is significantly different. For the first part of the study, the overall p-value was statistically significant (<0.0001 and $F = 16.39$). The analysis of the second part of the study, which compared different dose levels for the same thickness group, was performed considering each thickness range separately. The p-values were 0.0004 for T3 (50–69 mm) and 0.1305 for T4 (>70 mm).

Figure 4 shows an example of a simulated mass inserted in the four T-groups and acquired with standard AEC mode

TABLE I. Shown is the number of the true positive marks for all readers for each thickness group for the total of the lesions and for microcalcification clusters and masses separated. The total number of lesions, microcalcification clusters, and masses is also shown together with the percentage ratio. Results are shown for both parts of the study.

	Part 1 (standard AEC setting)			Part 2 (new AEC setting)		
	Correct lesions marked/total lesions	Correct microcalcification clusters marked/total microcalcification clusters	Correct masses marked/total masses	Correct lesions marked/total lesions	Correct microcalcification clusters marked/total microcalcification clusters	Correct masses marked/total masses
T1 (≤29 mm)	197/280 (70%)	86/140 (61%)	111/140 (79%)			
T2 (30–49 mm)	144/280 (50%)	81/140 (58%)	63/140 (45%)			
T3 (50–69 mm)	94/280 (34%)	45/140 (32%)	49/140 (35%)	130/280 (46%)	69/140 (49%)	61/140 (44%)
T4 (≥70 mm)	104/280 (37%)	37/140 (26%)	67/140 (48%)	121/280 (43%)	58/140 (41%)	63/140 (45%)

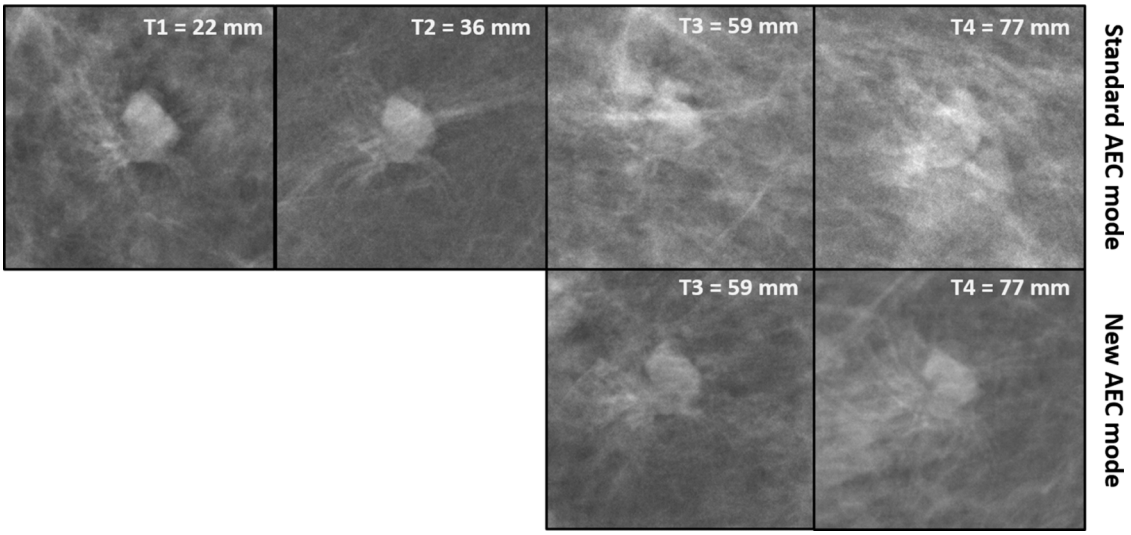

FIG. 4. A mass template inserted in (left column) the four T-groups, for standard AEC dose settings and (right column) T3 and T4 for the new AEC dose settings. All images were BI-RADS category 2 and the template was inserted in an area of VOLPARA glandularity value equal to 10%.

(left column) and acquired with the new AEC mode for T3 and T4 (right column). In this example, the readers did not detect the masses present in the images of T3 and T4 acquired with standard AEC dose while they successfully detected the masses of the same groups acquired with new AEC dose. A similar example for a microcalcification cluster is shown in Fig. 5.

3.C. Part 1—Lesion detectability as a function of breast thickness

3.C.1. Analysis for all lesions, for masses and for microcalcification clusters

The results of the FoM for part 1 of the study considering all lesions are shown in Fig. 6(a) and values are listed in Table III. The AFROC curves are plotted as lesion localization fraction (LLF) as a function of false positive fraction (FPF). The area under the AFROC curve decreases from 0.802 to 0.553 with increasing thickness from T1 to T3 while the AFROC area for T4 (0.565) was found to be almost equal to T3.

The results of the *p*-values and the 95% confidence interval (CI) for the paired comparison of the different thickness groups are presented in Table IV. A *p*-value smaller than 0.05 and a CI that does not include zero mean a significant difference between the two thickness groups being compared. All the *p*-values were smaller than 0.05 and the 95% CI did not include zero except for the comparison of T3 and T4 where the *p*-value was equal to 0.7757 and the 95% CI was (−0.0909,0.0680). This indicates that these two groups are considered not significantly different, which was expected given the similar values found for the AFROC areas of the T3 and T4 groups.

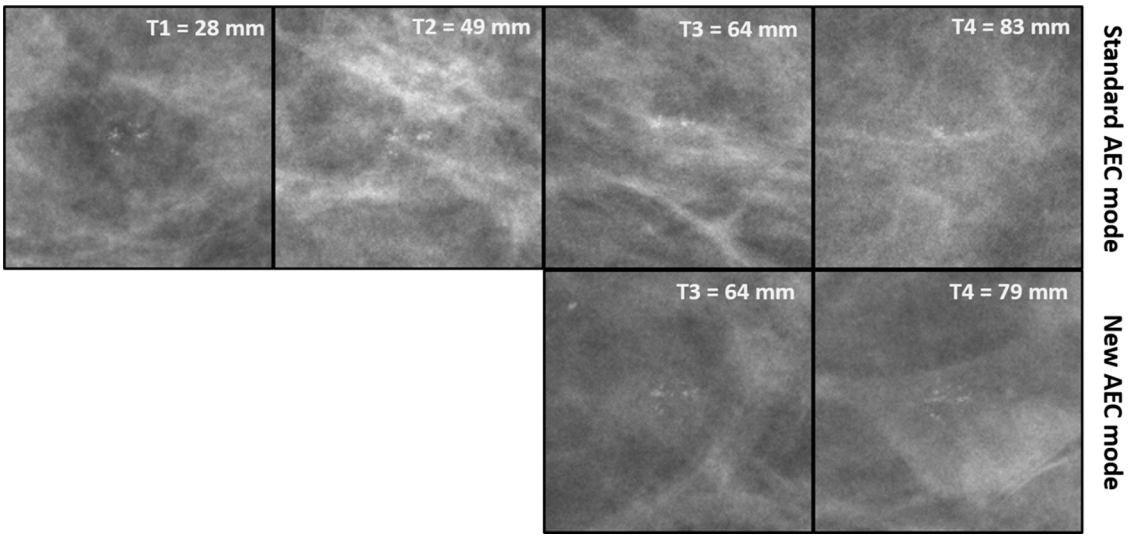

FIG. 5. The figure shows a microcalcification cluster template inserted in (right column) the four T-groups, for standard AEC dose settings and (left column) T3 and T4 for the new AEC dose settings All images were BI-RADS category 2 and the template was inserted in an area of VOLPARA glandularity value equal to 11%.

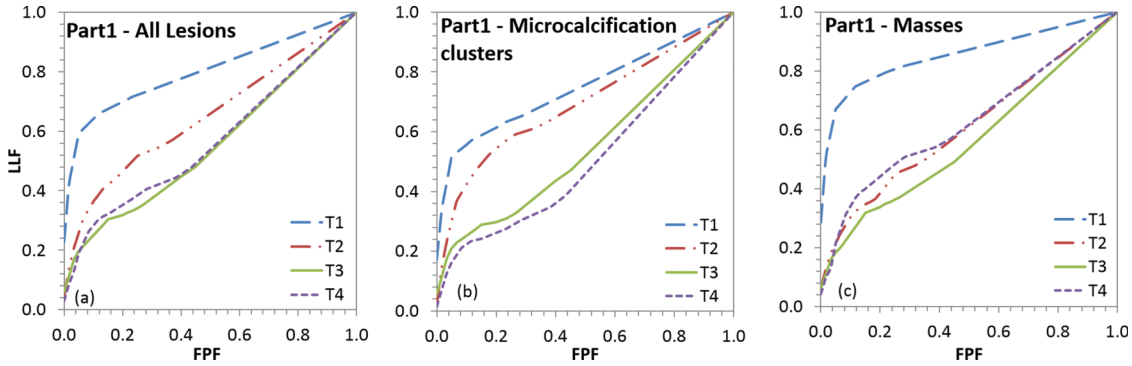

FIG. 6. AFROC curves for all readers for (a) all lesions (b) microcalcification clusters (c) masses. Results are for part 1 of the study for all thickness groups.

The same analysis performed for all lesions was repeated for microcalcification clusters and masses separately. The results found for microcalcifications, Fig. 6(b) and Table III, show a trend similar to the results for all lesions. The AFROC area falls from 0.749 to 0.506 as breast thickness increases. In Table IV, no significant differences were found for the paired comparison of T1–T2 and T3–T4. All of the other pairs of thickness groups were significantly different from each other. The analysis of the masses showed a different trend [Fig. 6(c) and Table III]. The T1 AFROC area is larger than the areas for the other T-groups - these remaining T-groups have a similar AFROC area, 0.609, 0.561 and 0.623 for T2, T3 and T4 respectively. The *p*-values listed in Table IV confirm these results showing *p*-values smaller than 0.05 when T1 is compared to all other groups while there is no significant difference for the other paired groups.

3.D. Part 2—Lesion detectability as a function of different AEC dose settings

3.D.1. Analysis for all lesions

Results for part 2 of the study, considering all lesions, are shown in Figs. 7(a) and 8(a) for T3 and T4 respectively. In these graphs, the AFROC curves are compared with the results obtained with the standard AEC setting (Fig. 6); FoM values of the curves are listed in Table V. The AFROC areas for both T3 and T4 are larger compared to the standard AEC setting used in part 1 and in particular they are both equal to 0.652. The results of the comparison of the paired groups for the different AEC settings, i.e., the comparison of part 1 to part 2, is shown in Table VI which shows a significant difference between the two settings for the thickness group T3 (*p*-value = 0.0004, 95% CI = 0.0443, 0.1531). There is no significant difference for the group T4 (*p*-values = 0.1305, 95% CI = −0.0337, 0.2076).

3.D.2. Analysis for microcalcification clusters and masses treated separately

The results for the separate analyses of microcalcification clusters and masses are similar to considering all lesions together. The AFROC areas increased for the new AEC setting when compared to the (lower dose) standard AEC. The results are shown in Figs. 7 and 8 and the values are listed in Table V. Considering only the microcalcification clusters, the paired settings for both groups T3 and T4 showed a significant difference between these groups. There is no significant difference for the analysis of the masses. The *p*-value for the T3 groups is equal to 0.0644 and for T4 0.5525.

3.E. Combined AEC settings

To conclude the study, an analysis of combined AEC settings was performed. This analysis evaluated the influence of thicknesses on detectability combining the results obtained in part 1 and 2. In particular the results of thickness groups T1 and T2 of part 1 (standard AEC dose settings) and the results of thickness groups T3 and T4 of part 2 (new AEC dose settings) were combined. The aim of this last analysis is to verify whether, by combining the settings, detectability of lesions can be considered constant for breast thicknesses above 30 mm. Results of the FoM, plotted as LLF as a function of FPF, are shown in Fig. 9 for (a) all lesions, (b)

TABLE III. Figure of merit for the JAFROC analysis (AFROC area) shown for the different thickness groups and investigated for the first part of the study. Results are presented for all lesions and for microcalcification clusters and masses separately.

FoM (AFROC area)—Study part 1 (standard AEC setting)			
Breast thickness groups compared	All lesions	Microcalcification clusters	Masses
T1 (≤29 mm)	0.802	0.749	0.855
T2 (30–49 mm)	0.650	0.691	0.609
T3 (50–69 mm)	0.553	0.545	0.561
T4 (≥70 mm)	0.565	0.506	0.623

TABLE IV. The p-values and 95% confidence interval for the comparison of the different breast thickness groups investigated in part 1 of the study. A p-value smaller than 0.05 and a CI that does not include zero mean a significant difference between the two investigated thickness groups. The symbol ˙ indicates no significant difference between the paired groups. Results are listed for all lesions and for microcalcification clusters and masses separately.

Study part 1 (standard AEC dose setting)						
Breast thickness groups compared	All lesions		Microcalcification clusters		Masses	
	p-value	95% CI	p-value	95% CI	p-value	95% CI
T1–T2	0.0003	0.0722, 0.2311	0.2322˙	−0.0377, 0.1519	0.0000	0.1360, 0.3565
T1–T3	0.0000	0.1694, 0.3284	0.0001	0.1085, 0.2982	0.0000	0.1842, 0.4047
T1–T4	0.0000	0.1580, 0.3169	0.0000	0.1480, 0.3376	0.0000	0.1218, 0.3423
T2–T3	0.0169	0.0177, 0.1767	0.0032	0.0514, 0.2411	0.3905˙	−0.0621, 0.1585
T2–T4	0.0346	0.0063, 0.1652	0.0003	0.0909, 0.2806	0.7999˙	−0.1245, 0.0960
T3–T4	0.7757˙	−0.0909, 0.0680˙	0.4064˙	−0.0553, 0.1344˙	0.2663˙	−0.1727, 0.0478˙

microcalcification clusters, and (c) masses. The values of the AFROC areas were previously listed in Tables III and V. For the analysis of all lesions in Fig. 9(a), it can be seen that there is almost no difference for the curves of the thickness groups T2, T3, and T4 with areas of 0.650, 0.652, and 0.652, respectively, while T1 has a larger area equal to 0.802. These results are confirmed by the p-values and 95% CI data listed in Table VII. Significant differences were obtained when the group T1 is compared to all the other groups, while there is no significant difference for all the other paired groups. The same trend was found for the analysis of the masses, Fig. 9(c) and Table VII. A slightly different trend is seen for the microcalcification clusters where the AFROC areas are equal to 0.749, 0.691, 0.665, and 0.643 going from T1 to T4. The p-values, Table VII, showed no significant difference between all paired groups, except for the comparison of T1 and T4 which has a p-value of 0.0408 and a 95% CI of (0.0863, 0.3042).

4. DISCUSSION

In the first part of this study the effect of compressed breast thickness on lesion detection was investigated using simulated lesions inserted in real mammograms acquired under the current standard clinical AEC settings. This is an AEC setup that produces a constant pixel value when tested with different thicknesses of homogenous PMMA plates, used as a surrogate for breast thickness. Using a test object as simple as a 2 mm thick Al sheet of 1×1 cm on a homogenous background as well as more sophisticated d' measures or contrast detail analyses show a reduction in performance with increasing PMMA thickness.[16] Results from the more detailed clinical simulation study reported here, showed a similar, significant decrease in lesion detection when compressed breast thickness increases. A similar decreasing trend was seen when the analysis was performed only for the microcalcification clusters while for the detection of masses this trend was less clear. The detection of masses for this AEC setting was not significantly different for compressed breast thicknesses above 30 mm [see Fig. 6(c)]. This reduction in detectability can be partly explained by the increasing quantity of scattered radiation, together with a reduction in primary contrast as breast thickness increases.[38,39] A number of studies have shown that the detection of objects smaller than 1 mm is mainly quantum noise limited while the detection of larger detail structures is mainly limited by breast parenchymal patterns.[40,41] This study uses real clinical images, with both quantum and anatomical structured noise levels set by the acquired mammography images. The absence of a drop in detectability of the masses in the highest thickness group may be due to features in the anatomical background for this

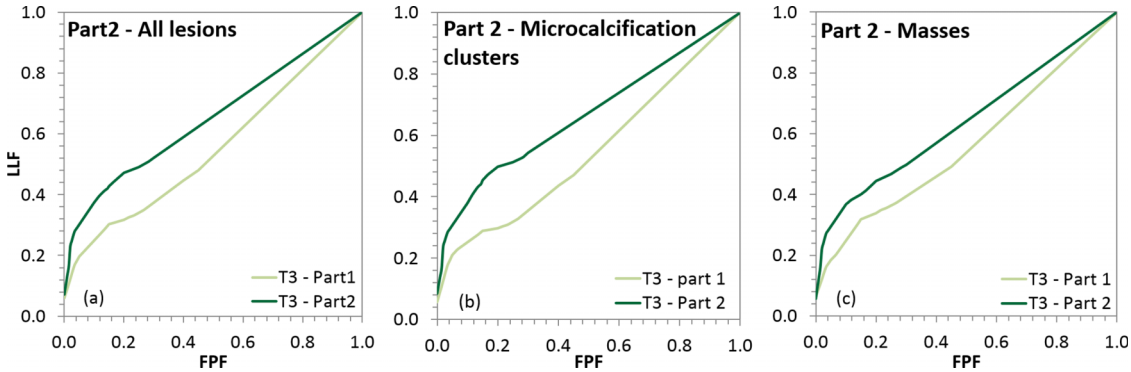

FIG. 7. AFROC curves for all readers for thickness group T3 (50–69 mm). (a) All lesions, (b) microcalcification clusters, and (c) masses. Comparison of the two different AEC dose settings (part 1: standard AEC dose, part 2: new AEC dose settings).

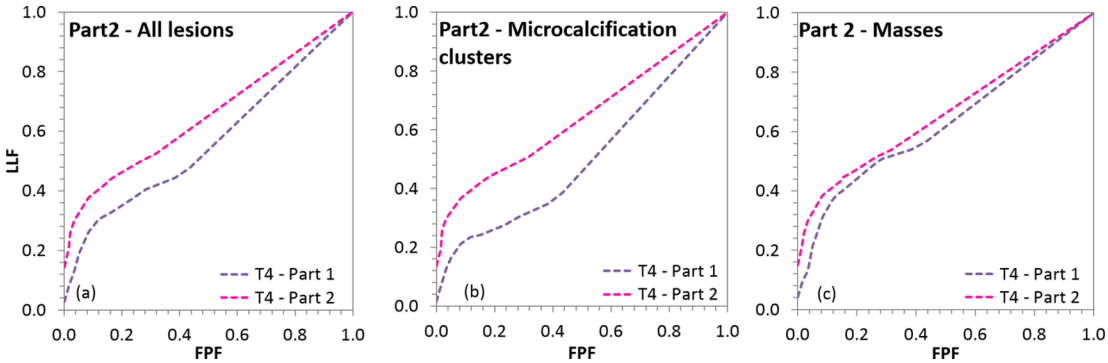

FIG. 8. Alternative free-response receiver operating characteristic curves (AFROC) for all readers for thickness group T4 (>70 mm). (a) All lesions, (b) microcalcification clusters, and (c) masses. Comparison of the two different AEC settings (part 1: standard AEC dose, part 2: new AEC dose settings).

thickness class. Although we have inserted each particular lesion model in a breast segment of a given glandularity measure that was fixed among the thickness groups, some particular aspects of the background structure could not be controlled. A typical example is the fact that the thickest breasts tend to have small islands of glandular tissue when compared to the more global distribution of glandular tissue in thinner breasts that tend to isolate the potential location of the simulated lesion and hence may influence mass detectability. This potential bias was not considered a negative aspect of the study as we aimed to be representative of the clinical environment.

In the second part of the work the mammograms used for insertion of simulated lesions were acquired with higher doses than the standard settings. These settings were based on the results of a previous study[17] describing an AEC setup that maintains a theoretical index of detectability constant as a function of object thickness. The results of the present study showed a significant increase in detection of all lesions for the thickness group T3 but not for T4. A significant increase was found in microcalcification cluster detection for both thickness groups (T3 and T4) while for masses the increase was not significant. Even if there is a lack of statistical significance, it is important to note that for the group T3 the detection rate of subtle masses has increased from 33% to 44%. Contrary to the result for microcalcifications, the number of detected masses decreased slightly between parts 1 and 2 for thickness group T4, resulting in a nonsignificant p-value ($p = 0.5525$) for the AFROC FoM. The AFROC curve itself for part 2 (T4) lies above that of part 1 (T4) for masses [Fig. 8(c)] and this is a result of different FP rates for these datasets. The number of nonvisible and obvious cases scored was not different between part 1 and part 2. A relevant further investigation could be the quantification of the number of subtle and obvious lesions detected in screening programs. These data would allow us to calculate the impact our study would have on a screening population. These results for microcalcifications and masses were previously seen in the work of Samei *et al.* and Ruschin *et al.*[14,15] These studies both investigated the detectability of microcalcification and masses at reduced dose levels by adding noise to images, and found a significant decrease in microcalcification detectability but no significant difference in mass detectability as a function of dose. In both studies, the authors suggested that any dose reduction should be applied with caution. For ethical reasons, most previous studies on mammography operating points have investigated the effects of dose reduction by simulating an increase in image noise. In fact, the study described here has taken the opposite approach, to increase the dose while remaining within both the current Achievable and Acceptable dose levels in the European Guidelines.[20] It is less clear cut that an increase in dose may lead to an increase in object detectability as structure noise may start to dominate object detection at higher doses. Figure 6(b) shows that structure noise is already dominating detection of masses even for the standard AEC setting.

This study has also shown that combining the standard mammographic settings for compressed breast thicknesses up to 49 mm with the increased doses settings for compressed breast thicknesses above 50 mm resulted in equal detectability for all lesions and masses for thicknesses above 30 mm. For microcalcification clusters, no significant difference between all T-groups, apart from the T1–T4 comparison was found. As a general conclusion, these data suggest that (technical) detectability performance can be set using test object methods as a function of simulated breast thickness and this performance follows through to clinically realistic detection results. In this work the detectability level was set at the level of the 30–49 mm group for all thick breasts. This compensated for the fall in detectability seen in constant PV AEC regimes: the situation of a decrease in detectability as a function of thickness was changed into an equal/constant detectability. This result was predicted in earlier studies[17,42] and therefore this study has confirmed the possibility of reaching constant lesion detectability in clinical images by reprogramming the AEC device, such

TABLE V. Figure of merit for the JAFROC analysis (AFROC area) shown for the different thickness groups for both parts of the study with microcalcification clusters and masses considered together and separately.

FoM (AFROC area)—Study part 2 (new AEC dose settings)			
	All lesions	Microcalcification clusters	Masses
T3 (50–69 mm)	0.652	0.665	0.638
T4 (≥70 mm)	0.652	0.643	0.660

TABLE VI. The table shows the p-values and the 95% confidence interval for the comparison of the different AEC settings investigated in part 2 of the study for the breast thickness groups T3 and T4. A p-value smaller than 0.05 and a CI that does not include zero means a significant difference between the two AEC settings investigated for a specific breast thickness group. The symbol ˙ indicates no significant difference. Results are listed for all lesions and for microcalcification clusters and masses separately.

Study part 2 (new AEC dose settings)						
AEC settings compared	All lesions		Microcalcification clusters		Masses	
	p-value	95% CI	p-value	95% CI	p-value	95% CI
T3–T3	0.0004	0.0443, 0.1531	0.0074	0.0364, 0.2037	0.0644˙	−0.0047, 0.1595
T4–T4	0.1305˙	−0.0337, 0.2076	0.0322	0.0148, 0.2594	0.5525˙	−0.0947, 0.1683

that a specific theoretical detectability index is kept constant. Moreover it has shown that conclusions derived from simple test objects, imaged against a homogeneous background, can make predictions of lesion detectability that apply to clinical images, containing structured anatomical noise. This finding has important implications for quality control and assurance programs, implying that that the methods applied by medical physicists can have a substantial, positive influence on clinical image quality. The constant PV AEC scheme is not just confined to the Siemens Inspiration systems used in this study but is currently implemented on most other mammography systems. We therefore expect a fall in lesion detectability with breast thickness for other such AEC designs and the modified AEC scheme to improve detectability for thicker breasts.

Increasing the doses in groups T3 and T4 appears to be a straightforward means of making lesion detectability more constant among women and significantly increasing lesion detection rates, but these increased doses and associated patient risk should be justified. Risk benefit calculations for breast cancer screening were considered beyond the scope of present work. In general, the increase of doses can be justified if this action results in an increase of lives saved by an earlier detection of cancer, balanced against cancers induced by the exposure. As a first approach, we can use the data of Law *et al.*[43] for the calculation of the ratio of detected/induced cancers in screen/film mammography. It can be seen that the doses used in this study for the modified AEC (part 2) are similar or lower to those used in screen/film mammography described by Law *et al.*[44] As a first step, we suggest that if benefit compared to the risk was acceptable for screen-film mammography, we can assume that it is also acceptable for the newly proposed AEC setting. In our study, we showed improved detectability for subtle lesions. Our radiologists stressed this fact: this feature will help clinicians find the smallest cancers that have a large impact on the efficiency of breast screening programs. The benefit to risk balance may therefore even be improved when compared to the standard AEC settings, but this certainly warrant further study.

Four radiologists (two from Belgium, one from Austria and one from the United Kingdom) were consulted about an increase of dose for breast thicknesses above 49 mm in order to obtain a more constant lesion detectability. This was done to gauge the willingness of clinicians to accept the use of increased doses for some patient groups. They all underlined the difficulty of imaging thick breasts, notwithstanding the general consensus regarding the easier reading of fatty breasts. It was felt that increasing dose has a significant effect on microcalcification cluster detection and there is a trend of improvement on the visibility of masses but when the difficulties related to the positioning and compression of thicker breast are taken into account, then the increase of dose could help to counteract some of these difficulties, leading to an improvement in lesion detection. In fact, for thick breasts, the combination of low doses along with nonoptimal breast compression breast could lead to even worse situations than those examined in part 1 of the study. Based on our results, radiologists concluded that the increase in dose is perhaps not needed for the visualization of small invasive carcinoma, provided all other aspects of image acquisition (i.e., compression and positioning of the

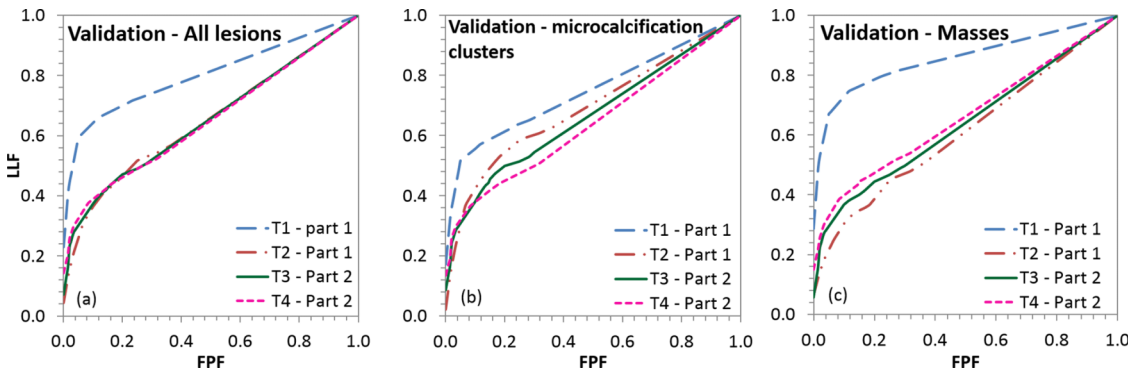

FIG. 9. Alternative free-response receiver operating characteristic curves (AFROC) for all readers for (a) all lesions, (b) microcalcification clusters, and (c) masses. Results are for parts 1 and 2 of the study combined for all thickness groups.

TABLE VII. The table shows the p-values and the 95% confidence interval for the comparison of thickness groups T1 and T2 of part 1 and T3 and T4 of part 2. A p-value smaller than 0.05 and a CI that does not include zero means a significant difference between the two AEC settings investigated for a specific breast thickness group. The symbol ˙ indicates no significant difference. Results are listed for all lesions and for microcalcification clusters and masses separately.

Combined AEC settings						
Breast thickness groups compared	All lesions		Microcalcification clusters		Masses	
	p-value	95% CI	p-value	95% CI	p-value	95% CI
T1–T2	0.0005	0.0692, 0.2342	0.2602˙	−0.0440, 0.1582˙	0.0000	0.1373, 0.3552
T1–T3	0.0006	0.0677, 0.2327	0.1036˙	−0.0178, 0.1844˙	0.0001	0.1081, 0.3560
T1–T4	0.0006	0.0680, 0.2330	0.0408	0.0046, 0.2068	0.0005	0.0863, 0.3042
T2–T3	0.9713˙	−0.0840, 0.0810˙	0.6031˙	−0.0749, 0.1273˙	0.5982˙	−0.1380, 0.0798
T2–T4	0.9771˙	−0.0837, 0.0813˙	0.3364˙	−0.0524, 0.1497˙	0.3570˙	−0.1599, 0.0579
T3–T4	0.9943˙	−0.0822, 0.0828˙	0.6558˙	−0.0786, 0.1235˙	0.6930˙	−0.1308, 0.0871˙

breast) are optimized. On the other hand, the significant decrease of visibility for the microcalcification clusters using the standard AEC settings will mean that the first signs of malignancy go largely undetected in larger breasts. The radiologists interviewed in the frame of the present study indicated a definite preference for the new AEC settings. The solution to improving sensitivity and specificity of lesion detection in screening programs will be a combination of factors, including the use of optimized AEC regimes together with the use of advanced 3D breast imaging techniques such as breast tomosynthesis, for lesions of which detection is currently limited by breast anatomical structure.

5. CONCLUSIONS

This work has investigated lesion detectability as a function of compressed breast thickness via the simulation of masses and microcalcification clusters into real mammograms. Detectability for simulated lesions was found to decrease as compressed breast thickness increased, as predicted by technical Quality Control measurements. A modified AEC setup was implemented for breast thicknesses above 50 mm, in which dose to the detector was varied in response to increasing breast thickness in order to maintain technical (QC) measures of detectability constant. The modified AEC scheme increased detectability significantly for microcalcifications simulated into the two largest thickness groups, bringing detectability in breasts of thickness 30 mm and above up to a constant level. For the group T4 (breast thickness ≥70 mm), the dose increase did not yield an improvement in mass lesion detectability, suggesting that alternative methods such as breast tomosynthesis are needed to overcome anatomical structure noise and improve mass detection. These results reinforce the utility of technical measurements made as part of routine QC performance assessments and that such data are in fact related to the detection of simulated lesions in real clinical backgrounds. Implementation of the modified AEC scheme should help to move 2D mammography imaging towards constant lesion detectability, irrespective of compressed breast thickness.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Eman Shaheen for providing the simulation software and the lesion models. The authors thank the radiologists for their time spent in reading the hybrid images: Professor Dr. Chantal Van Ongeval, Dr. Julie Soens, Dr. Mieke Smets, and Dr. Liesbeth Cornelissen. The authors would also like to acknowledge Professor Dr. D. P. Chakraborty for all prompt replies regarding their questions on statistics. Thanks to the VOLPARA team, in particular to the authors' local application specialists Rob Maas and Tom Van Laer, for providing the software, for help with the installation and for technical support. Finally the authors want to acknowledge Thomas Mertelmeier, Anna Jerebko, and Andreas Fieselmann from Siemens for helping in the setup of the AEC device and for providing the software to apply the clinical image processing.

CONFLICT OF INTEREST DISCLOSURE

The authors have no COI to report.

[a]Author to whom correspondence should be addressed. Electronic mail: elena.salvagnini@gmail.com

[1]A. M. J. Bluekens, R. Holland, N. Karssemeijer, M. J. M. Broeders, and G. J. den Heeten, "Comparison of digital screening mammography and screen-film mammography in the early detection of clinically relevant cancers: A multicenter study," Radiology **265**, 707–714 (2012).
[2]J. M. Lewin, R. E. Hendrick, C. J. D'Orsi, P. K. Isaacs, L. J. Moss, A. Karellas, G. A. Sisney, C. C. Kuni, and G. R. Cutter, "Comparison of full-field digital mammography with screen-film mammography for cancer detection: Results of 4945 paired examinations," Radiology **218**(3), 873–880 (2001).
[3]P. Skaane, C. Balleyguier, F. Diekmann, S. Diekmann, J. C. Piguet, K. Young, and L. T. Niklason, "Breast lesion detection and classification: Comparison of screen-film mammography and full-field digital mammography with soft-copy reading—Observer performance study," Radiology **237**(1), 37–44 (2005).
[4]P. Skaane, K. Young, and A. Skjennald, "Population-based mammography screening: Comparison of screen-film and full-field digital mammography with soft-copy reading—Oslo I study," Radiology **229**(3), 877–884 (2003).
[5]N. M. Hambly, M. M. McNicholas, N. Phelan, G. C. Hargaden, A. O'Doherty, and F. L. Flanagan, "Comparison of digital mammography and screen-film mammography in breast cancer screening: A review in the Irish breast screening program," Am. J. Roentgenol. **193**(4), 1010–1018 (2009).

[6]A. K. Bloomquist, M. J. Yaffe, E. D. Pisano, R. E. Hendrick, G. E. Mawdsley, S. Bright, S. Z. Shen, M. Mahesh, E. L. Nickoloff, R. C. Fleischman, M. B. Williams, A. D. Maidment, D. J. Beideck, J. Och, and J. A. Seibert, "Quality control for digital mammography in the ACRIN DMIST trial: Part I," Med. Phys. **33**(3), 719–736 (2006).
[7]M. J. Yaffe, A. K. Bloomquist, G. E. Mawdsley, E. D. Pisano, R. E. Hendrick, L. L. Fajardo, J. M. Boone, K. Kanal, M. Mahesh, R. O. Fleischman, J. Och, M. B. Williams, D. J. Beideck, and A. D. Maidment, "Quality control for digital mammography: Part II recommendations from the ACRIN DMIST trial," Med. Phys. **33**(3), 737–752 (2006).
[8]D. R. Dance, A. Thilander Klang, M. Sandborg, C. L. Skinner, S. A. Castellano, and G. Alm Carlsson, "Influence of anode/filter material and tube potential on contrast, signal-to-noise ratio and average absorbed dose in mammography: A Monte Carlo study," Br. J. Radiol. **73**(874), 1056–1067 (2000).
[9]P. Toroi, F. Zanca, K. C. Young, C. van Ongeval, G. Marchal, and H. Bosmans, "Experimental investigation on the choice of the tungsten/rhodium anode/filter combination for an amorphous selenium-based digital mammography system," Eur. Radiol. **17**(9), 2368–2375 (2007).
[10]P. Bernhardt, T. Mertelmeier, and M. Hoheisel, "X-ray spectrum optimization of full-field digital mammography: Simulation and phantom study," Med. Phys. **33**(11), 4337–4349 (2006).
[11]M. B. Williams, P. Raghunathan, M. J. More, J. A. Seibert, A. Kwan, J. Y. Lo, E. Samei, N. T. Ranger, L. L. Fajardo, A. McGruder, S. M. McGruder, A. D. Maidment, M. J. Yaffe, A. Bloomquist, and G. E. Mawdsley, "Optimization of exposure parameters in full field digital mammography," Med. Phys. **35**(6), 2414–2423 (2008).
[12]N. W. Marshall, "Detective quantum efficiency measured as a function of energy for two full-field digital mammography systems," Phys. Med. Biol. **54**(9), 2845–2861 (2009).
[13]E. Salvagnini, H. Bosmans, L. Struelens, and N. W. Marshall, "Effective detective quantum efficiency for two mammography systems: Measurement and comparison against established metrics," Med. Phys. **40**(10), 101916 (16pp.) (2013).
[14]E. Samei, R. S. Saunders, J. A. Baker, and D. M. Delong, "Digital mammography: Effects of reduced radiation dose on diagnostic performance," Radiology **243**(2), 396–404 (2007).
[15]M. Ruschin, P. Timberg, M. Båth, B. Hemdal, T. Svahn, R. S. Saunders, E. Samei, I. Andersson, S. Mattsson, D. P. Chakrabort, and A. Tingber, "Dose dependence of mass and microcalcification detection in digital mammography: Free response human observer studies," Med. Phys. **34**(2), 400–407 (2007).
[16]E. Salvagnini, H. Bosmans, L. Struelens, and N. W. Marshall, "Tailoring automatic exposure control toward constant detectability in digital mammography," Med. Phys. **42**(7), 3834–3847 (2015).
[17]E. Salvagnini, H. Bosmans, P. Monnin, L. Struelens, F. R. Verdun, and N. W. Marshall, "The use of detectability indices as a means of automatic exposure control for a digital mammography system," Proc. SPIE **7961**, 79615J (2011).
[18]N. Van Peteghem, E. Salvagnini, H. Bosmans, L. Cockmartin, and N. W. Marshall, "A comparison of mammographic systems for different breast thicknesses using model observer detectability," Proc. SPIE **9412**, 941235 (2015).
[19]E. Salvagnini, H. Bosmans, L. Struelens, and N. W. Marshall, "Effective detective quantum efficiency (eDQE) and effective noise equivalent quanta (eNEQ) for system optimization purposes in digital mammography," Proc. SPIE **8313**, 83130H (2012).
[20]European Commission, "European guidelines for breast cancer screening," in *The European Protocol for the Quality Control of the Physical and Technical Aspects of Mammography Screening. Part B: Digital Mammography*, 4th ed. (European Commission, Luxembourg, 2006).
[21]R. Highnam, S. M. Brady, M. Yaffe, N. Karssemeijer, and J. Harvey, "Robust breast composition measurement—VOLPARA™," in *Proceedings of IWDM*, 2010.
[22]A. Gubern-Mérida, M. Kallenberg, B. Platel, R. M. Mann, R. Martí, and N. Karssemeijer, "Volumetric breast density estimation from full-field digital mammograms: A validation study," PLoS One **9**(1), e85952 (2014).
[23]S. van Engeland, P. R. Snoeren, H. Huisman, C. Boetes, and N. Karssemeijer, "Volumetric breast density estimation from full-field digital mammograms," IEEE Trans. Med. Imaging **25**(3), 273–282 (2006).
[24]E. Shaheen, F. De Keyzer, H. Bosmans, D. R. Dance, K. C. Young, and C. Van Ongeval, "The simulation of 3D mass models in 2D digital mammography and breast tomosynthesis," Med. Phys. **41**(8), 081913 (8pp.) (2014).
[25]E. Shaheen, C. Van Ongeval, F. Zanca, L. Cockmartin, N. W. Marshall, J. Jacobs, K. C. Young, D. R. Dance, and H. Bosmans, "The simulation of 3D microcalcification clusters in 2D digital mammography and breast tomosynthesis," Med. Phys. **38**(12), 6659–6671 (2011).
[26]F. Zanca, J. Jacobs, C. Van Ongeval, F. Claus, V. Celis, C. Geniets, V. Provost, H. Pauwels, G. Marchal, and H. Bosmans, "Evaluation of clinical image processing algorithms used in digital mammography," Med. Phys. **36**(3), 765–775 (2009).
[27]L. M. Warren, R. M. Given-Wilson, M. G. Wallis, J. Cooke, M. D. Halling-Brown, A. Mackenzie, D. P. Chakraborty, H. Bosmans, D. R. Dance, and K. C. Young, "The effect of image processing on the detection of cancers in digital mammography," AJR, Am. J. Roentgenol. **203**(2), 387–393 (2014).
[28]E. Shaheen, F. Zanca, F. Sisini, G. Zhang, J. Jacobs, and H. Bosmans, "Simulation of 3D objects into breast tomosynthesis images," Radiat. Prot. Dosim. **139**(1-3), 108–112 (2010).
[29]E. Salvagnini, H. Bosmans, L. Struelens, and N. W. Marshall, "Quantification of scattered radiation in projection mammography: Four practical methods compared," Med. Phys. **39**(6), 3167–3180 (2012).
[30]D. P. Chakraborty and K. S. Berbaum, "Observer studies involving detection and localization: Modeling, analysis, and validation," Med. Phys. **31**(8), 2313–2330 (2004).
[31]D. P. Chakraborty, "A brief history of free-response receiver operating characteristic paradigm data analysis," Acad. Radiol. **20**(7), 915–919 (2013).
[32]S. Börjesson, M. Håkansson, M. Båth, S. Kheddache, S. Svensson, A. Tingberg, A. Grahn, M. Ruschin, B. Hemdal, S. Mattsson, and L. G. Månsson, "A software tool for increased efficiency in observer performance studies in radiology," Radiat. Prot. Dosim. **114**(1-3), 45–52 (2005).
[33]M. Håkansson, S. Svensson, S. Zachrisson, A. Svalkvist, M. Båth, and L. G. Månsson, "VIEWDEX: An efficient and easy-to-use software for observer performance studies," Radiat. Prot. Dosim. **139**(1-3), 42–51 (2010).
[34]M. Håkansson, S. Svensson, S. Zachrisson, A. Svalkvist, M. Båth, and L. G. Månsson, "On the choice of acceptance radius in free-response observer performance studies," Br. J. Radiol. **86**(1021), 42313554 (2013).
[35]D. D. Dorfman, K. S. Berbaum, and C. E. Metz, "Receiver operating characteristic rating analysis: Generalization to the population of readers and patients with the jackknife method," Invest. Radiol. **27**, 723–731 (1992).
[36]S. L. Hillis, K. S. Berbaum, and C. E. Metz, "Recent developments in the Dorfman-Berbaum-Metz procedure for multireader ROC study analysis," Acad. Radiol. **15**(5), 647–661 (2008).
[37]E. A. Berns, R. E. Hendrick, and G. R. Cutter, "Performance comparison of full-field digital mammography to screen-film mammography in clinical practice," Med. Phys. **29**(5), 830–834 (2002).
[38]U. Neitzel, "Grids or air gaps for scatter reduction in digital radiography: A model calculation," Med. Phys. **19**(2), 475–481 (1992).
[39]R. F. Wagner, G. T. Barnes, and B. S. Askins, "Effect of reduced scatter on radiographic information content and patient exposure: A quantitative demonstration," Med. Phys. **7**(1), 13–18 (1980).
[40]C. J. Kotre, "The effect of background structure on the detection of low contrast objects in mammography," Br. J. Radiol. **71**(851), 1162–1167 (1998).
[41]W. Huda, K. M. Ogden, E. M. Scalzetti, D. R. Dance, and E. A. Bertrand, "How do lesion size and random noise affect detection performance in digital mammography?," Acad. Radiol. **13**(11), 1355–1366 (2006).
[42]I. M. Suarez Castellanos, R. Kaczmarek, C. C. Brunner, H. de Las Heras, H. Liu, and K. Chakrabarti, "Evaluation of automatic exposure control performance in full-field digital mammography systems using contrast-detail analysis," Proc. SPIE **8313**, 83134I (2012).
[43]J. Law, K. Faulkner, and K. C. Young, "Risk factors for induction of breast cancer by x-rays and their implications for breast screening," Br. J. Radiol. **80**(952), 261–266 (2007).
[44]K. Faulkner, "Mammographic screening: Is the benefit worth the risk?," Radiat. Prot. Dosim. **117**(1-3), 318–320 (2006).